Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC


DOI
10.1016/j.physletb.2012.08.020

Publication date
2012

Document Version
Final published version

Published in
Physics Letters B
Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC

ATLAS Collaboration

This paper is dedicated to the memory of our ATLAS colleagues who did not live to see the full impact and significance of their contributions to the experiment.

1. Introduction

The Standard Model (SM) of particle physics [1–4] has been tested by many experiments over the last four decades and has been shown to successfully describe high energy particle interactions. However, the mechanism that breaks electroweak symmetry in the SM has not been verified experimentally. This mechanism [5–10], which gives mass to massive elementary particles, implies the existence of a scalar particle, the SM Higgs boson. The search for the Higgs boson, the only elementary particle in the SM that has not yet been observed, is one of the highlights of the Large Hadron Collider [11] (LHC) physics programme.

Indirect limits on the SM Higgs boson mass of \( m_H < 158 \text{ GeV} \) at 95% confidence level (CL) have been set using global fits to precision electroweak results [12]. Direct searches at LEP [13], the Tevatron [14–16] and the LHC [17,18] have previously excluded, at 95% CL, a SM Higgs boson with mass below 600 GeV, apart from some mass regions between 116 GeV and 127 GeV.

Both the ATLAS and CMS Collaborations reported excesses of events in their 2011 datasets of proton–proton (pp) collisions at centre-of-mass energy \( \sqrt{s} = 7 \text{ TeV} \) at the LHC, which were compatible with SM Higgs boson production and decay in the mass region 124–126 GeV, with significances of 2.9 and 3.1 standard deviations (σ), respectively [17,18]. The CDF and DØ experiments at the Tevatron have also recently reported a broad excess in the mass region 124–126 GeV, with significances of 2.9 and 3.1 standard deviations (σ) [14–16].

The previous ATLAS searches in 4.6–4.8 fb\(^{-1}\) of data at \( \sqrt{s} = 7 \text{ TeV} \) are combined here with new searches for \( H \to ZZ^{(*)} \to 4\ell \), \( H \to WW^{(*)} \to e\nu\nu \) and \( H \to WW^{(*)} \to \tau^+\tau^- \) in the 5.8–5.9 fb\(^{-1}\) of pp collision data taken at \( \sqrt{s} = 8 \text{ TeV} \) between April and June 2012.

The data were recorded with instantaneous luminosities up to \( 6.8 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1} \); they are therefore affected by multiple pp collisions occurring in the same or neighbouring bunch crossings (pile-up). In the 7 TeV data, the average number of interactions per bunch crossing was approximately 10; the average increased to approximately 20 in the 8 TeV data. The reconstruction, identification and isolation criteria used for electrons and photons in the 8 TeV data are improved, making the \( H \to ZZ^{(*)} \to 4\ell \) and \( H \to WW^{(*)} \to e\nu\nu \) searches more robust against the increased pile-up. These analyses were re-optimised with simulation and frozen before looking at the 8 TeV data.

In the \( H \to WW^{(*)} \to e\nu\nu \) channel, the increased pile-up deteriorates the event missing transverse momentum, \( E_{T,\text{miss}} \), resolution, which results in significantly larger Drell–Yan background in the same-flavour final states. Since the \( e\nu \) channel provides most of the sensitivity of the search, only this final state is used in the analysis of the 8 TeV data. The kinematic region in which a SM Higgs boson with a mass between 110 GeV and 140 GeV is observed is 120–135 GeV; using the existing LHC constraints, the observed local significances for \( m_H = 125 \text{ GeV} \) are 2.7σ for CDF [14], 1.1σ for DØ [15] and 2.8σ for their combination [16].

1 The symbol \( \ell \) stands for electron or muon.
searched for was kept blinded during the analysis optimisation, until satisfactory agreement was found between the observed and predicted numbers of events in control samples dominated by the principal backgrounds.

This Letter is organised as follows. The ATLAS detector is briefly described in Section 2. The simulation samples and the signal predictions are presented in Section 3. The analyses of the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ channels are described in Sections 4–6, respectively. The statistical procedure used to analyse the results is summarised in Section 7. The systematic uncertainties which are correlated between datasets and search channels are described in Section 8. The results of the combination of all channels are reported in Section 9, while Section 10 provides the conclusions.

2. The ATLAS detector

The ATLAS detector [19–21] is a multipurpose particle physics apparatus with forward–backward symmetric cylindrical geometry. The inner tracking detector (ID) consists of a silicon pixel detector, a silicon microstrip detector (SCT), and a straw-tube transition radiation tracker (TRT). The ID is surrounded by a thin superconducting solenoid which provides a 2 T magnetic field, and by high-granularity liquid-argon (LAr) sampling electromagnetic calorimeter. The electromagnetic calorimeter is divided into a central barrel (pseudorapidity $|\eta| < 1.475$) and end-cap regions on either end of the detector ($1.375 < |\eta| < 2.5$ for the outer barrel and $2.5 < |\eta| < 3.2$ for the inner wheel). In the region matched to the ID ($|\eta| < 2.5$), it is radially segmented into three layers. The first layer has a fine segmentation in $\phi$ to facilitate $\gamma/\gamma$ separation from $\pi^{0}$ and to improve the resolution of the shower position and direction measurements. In the region $|\eta| < 1.8$, the electromagnetic calorimeter is preceded by a presampler detector to correct for upstream energy losses. An iron-scintillator/tile calorimeter gives hadronic coverage in the central rapidity range ($|\eta| < 1.7$), while a LAr hadronic end-cap calorimeter provides coverage over $1.5 < |\eta| < 3.2$. The forward regions ($3.2 < |\eta| < 4.9$) are instrumented with LAr calorimeters for both electromagnetic and hadronic measurements. The muon spectrometer (MS) surrounds the calorimeters and consists of three large air-core superconducting magnets providing a toroidal field, each with eight coils, a system of precision tracking chambers, and fast detectors for triggering. The combination of all these systems provides charged particle measurements together with efficient and precise lepton and photon measurements in the pseudorapidity range $|\eta| < 2.5$. Jets and $E_{T}^{miss}$ are reconstructed using energy deposits over the full coverage of the calorimeters, $|\eta| < 4.9$.

3. Signal and background simulation samples

The SM Higgs boson production processes considered in this analysis are the dominant gluon fusion ($gg \rightarrow H$, denoted $ggF$), vector-boson fusion ($qq' \rightarrow qH$, denoted VBF) and Higgs-strahlung ($qq' \rightarrow WH$, denoted $WH$), $ZH$, and $H \rightarrow t\bar{t}H$ pair ($qq'/gg \rightarrow t\bar{t}H$, denoted $t\bar{t}H$) is taken into account only in the $H \rightarrow \gamma\gamma$ analysis. For the $ggF$ process, the signal cross section is computed by taking into account only in the $H \rightarrow \gamma\gamma$ analysis.

Table 1 lists the signal processes included in the analysis. The SM Higgs boson production processes considered in this analysis are the dominant gluon fusion ($gg \rightarrow H$, denoted $ggF$), vector-boson fusion ($qq' \rightarrow qH$, denoted VBF) and Higgs-strahlung ($qq' \rightarrow WH$, denoted $WH$), $ZH$, and $H \rightarrow t\bar{t}H$ pair ($qq'/gg \rightarrow t\bar{t}H$, denoted $t\bar{t}H$) is taken into account only in the $H \rightarrow \gamma\gamma$ analysis. For the $ggF$ process, the signal cross section is computed by taking into account only in the $H \rightarrow \gamma\gamma$ analysis.
data are applied to account for small differences between data and simulation (e.g. large samples of $W$, $Z$ and $J/\psi$ decays are used to derive scale factors for lepton reconstruction and identification efficiencies).

4. $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel

The search for the SM Higgs boson through the decay $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$, where $\ell = e$ or $\mu$, provides good sensitivity over a wide mass range (110–600 GeV), largely due to the excellent momentum resolution of the ATLAS detector. This analysis searches for Higgs boson candidates by selecting two pairs of iso-

1. Event selection

The data are selected using single-lepton or dilepton triggers. For the single-muon trigger, the $p_T$ threshold is 18 GeV for the 7 TeV data and 24 GeV for the 8 TeV data, while for the single-electron trigger the transverse energy $E_T$, threshold varies from 20 GeV to 22 GeV for the 7 TeV data and is 24 GeV for the 8 TeV data. For the dielectron triggers, the thresholds are 12 GeV for both electrons. For the dimuon triggers, the thresholds for the 7 TeV data are 10 GeV for each muon, while for the 8 TeV data the thresholds are 13 GeV. An additional asymmetric dimuon trigger is used in the 8 TeV data with thresholds 18 GeV and 8 GeV for the leading and sub-leading muon, respectively.

Muon candidates are formed by matching reconstructed 1D tracks with either a complete track or a track-segment reconstructed in the MS [84]. The muon acceptance is extended with respect to Ref. [83] using tracks reconstructed in the forward region of the MS (2.5 < $|\eta|$ < 2.7), which is outside the ID coverage. If both an ID and a complete MS track are present, the two independent momentum measurements are combined; otherwise the information of the ID or the MS is used alone. Electron candidates must have a well-reconstructed 1D track pointing to an electromagnetic calorimeter cluster and the cluster should satisfy a set of identification criteria [85] that require the longitudinal and transverse shower profiles to be consistent with those expected for electromagnetic showers. Tracks associated with electromagnetic clusters are fitted using a Gaussian-Sum Filter [86], which allows for bremsstrahlung energy losses to be taken into account.

Each electron (muon) must satisfy $p_T > 7$ GeV ($p_T > 6$ GeV) and be measured in the pseudorapidity range $|\eta| < 2.47$ ($|\eta| < 2.7$). All possible quadruplet combinations with same-flavour opposite-charge lepton pairs are then formed. The most energetic lepton in the quadruplet must satisfy $p_T > 20$ GeV, and the second (third) lepton in $p_T$ order must satisfy $p_T > 15$ GeV ($p_T > 10$ GeV). At least one of the leptons must satisfy the single-lepton trigger or one pair must satisfy the dilepton trigger requirements. The leptons are required to be separated from each other by $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} > 0.1$ if they are of the same flavour and by $\Delta R > 0.2$ otherwise. The longitudinal impact parameters of the leptons along the beam axis are required to be within 10 mm of the reconstructed primary vertex. The primary vertex used for the event is defined as the reconstructed vertex with the highest $\sum p_T^2$ of associated tracks and is required to have at least three tracks with $p_T > 0.4$ GeV. To reject cosmic rays, muon tracks are required to have a transverse impact parameter, defined as the distance of closest approach to the primary vertex in the transverse plane, of less than 1 mm.

The same-flavour and opposite-charge lepton pair with an invariant mass closest to the $Z$ boson mass ($m_2$) in the quadruplet is referred to as the leading lepton pair. Its invariant mass, denoted by $m_{12}$, is required to be between 50 GeV and 106 GeV. The remaining same-flavour, opposite-charge lepton pair is the subleading lepton pair. Its invariant mass, $m_{34}$, is required to be in the range $m_{34} < m_{12} < 115$ GeV, where the value of $m_{34}$ depends on the reconstructed four-lepton invariant mass, $m_{4\ell}$. The value of $m_{34}$ varies monotonically from 17.5 GeV at $m_{4\ell} = 120$ GeV to 50 GeV at $m_{4\ell} = 190$ GeV [87] and is constant above this value. All possible lepton pairs in the quadruplet that have the same flavour and opposite charge must satisfy $m_{12} > 5$ GeV in order to reject backgrounds involving the production and decay of $J/\psi$ mesons. If two or more quadruplets satisfy the above selection, the one with the highest value of $m_{34}$ is selected. Four different analysis sub-channels, 4e, 2e2$\mu$, 2$\mu$2$\ell$ and 4$\mu$, arranged by the flavour of the leading lepton pair, are defined.

Non-prompt leptons from heavy flavour decays, electrons from photon conversions and jets mis-identified as electrons have broader transverse impact parameter distributions than prompt leptons from $Z$ boson decays and/or are non-isolated. Thus, the $Z$ + jets and $t\bar{t}$ background contributions are reduced by applying a cut on the transverse impact parameter significance, defined as the transverse impact parameter divided by its uncertainty, $d_0/\sigma_{d_0}$. This is required to be less than $3.5$ ($6.5$) for muons (electrons). The electron impact parameter is affected by bremsstrahlung and thus has a broader distribution.

In addition, leptons must satisfy isolation requirements based on tracking and calorimetric information. The normalised track isolation discriminant is defined as the sum of the transverse momenta of tracks inside a cone of size $\Delta R = 0.2$ around the lepton direction, excluding the lepton track, divided by the lepton $p_T$. The tracks considered in the sum are those compatible with the lepton vertex and have $p_T > 0.4$ GeV ($p_T > 1$ GeV) in the case of electron (muon) candidates. Each lepton is required to have a normalised track isolation smaller than 0.15. The normalised calorimetric isolation for electrons is computed as the sum of the $E_T$ of positive-energy topological clusters [88] with a reconstructed barycentre falling within a cone of size $\Delta R = 0.2$ around the candidate electron cluster, divided by the electron $E_T$. The algorithm for topological clustering suppresses noise by keeping cells with a significant energy deposit and their neighbours. The summed energy of the cells assigned to the electron cluster is excluded, while a correction is applied to account for the electron energy deposited outside the cluster. The ambient energy deposition in the event from pile-up and the underlying event is accounted for using a calculation of the median transverse energy density from low-$p_T$ jets [89,90]. The normalised calorimetric isolation for electrons is required to be less than 0.20. The normalised calorimetric isolation discriminant for muons is defined by the ratio to the $p_T$ of the muon of the $E_T$ sum of the calorimeter cells inside a cone of size
\[ \Delta R = 0.2 \] around the muon direction minus the energy deposited by the muon. Muons are required to have a normalised calorimetric isolation less than 0.30 (0.15 for muons without an associated ID track). For both the track- and calorimeter-based isolation, any contributions arising from other leptons of the quadruplet are subtracted.

The combined signal reconstruction and selection efficiencies for a SM Higgs with \( m_H = 125 \) GeV for the 7 TeV (8 TeV) data are 37% (36%) for the 4\( \mu \) channel, 20% (22%) for the 2e2\( \mu \)/2\( \mu 2\mu \) channels and 15% (20%) for the 4\( e \) channel.

The 4\( \ell \) invariant mass resolution is improved by applying a Z-mass constrained kinematic fit to the leading lepton pair for \( m_{4\ell} < 190 \) GeV and to both lepton pairs for higher masses. The expected width of the reconstructed mass distribution is dominated by the experimental resolution for \( m_H < 350 \) GeV, and by the natural width of the Higgs boson for higher masses (30 GeV at \( m_H = 400 \) GeV). The typical mass resolutions for \( m_H = 125 \) GeV are 1.7 GeV, 1.7 GeV/2.2 GeV and 2.3 GeV for the 4\( \mu \), 2e2\( \mu \)/2\( \mu 2\mu \) and 4e sub-channels, respectively.

### 4.2. Background estimation

The expected background yield and composition are estimated using the MC simulation normalised to the theoretical cross section for \( ZZ^{(*)} \) production and by methods using control regions from data for the \( Z + \text{jets and } t\bar{t} \) processes. Since the background composition depends on the flavour of the sub-leading lepton pair, different approaches are taken for the \( \ell\ell + \mu \mu \) and the \( \ell\ell + ee \) final states. The transfer factors needed to extrapolate the background yields from the control regions defined below to the signal region are obtained from the MC simulation. The MC description of the selection efficiencies for the different background components has been verified with data.

The reducible \( \ell\ell + \mu \mu \) background is dominated by \( t\bar{t} \) and \( Z + \text{jets (mostly } Zb \text{ ) events. A control region is defined by removing the isolation requirement on the leptons in the sub-leading pair, and by requiring that at least one of the sub-leading muons fails the transverse impact parameter significance selection. These modifications remove \( ZZ^{(*)} \) contributions, and allow both the \( t\bar{t} \) and \( Z + \text{jets backgrounds to be estimated simultaneously using a fit to the } m_{12} \text{ distribution. The } t\bar{t} \text{ background contribution is cross-checked by selecting a control sample of events with an opposite charge } e\mu \text{ pair with an invariant mass between 50 GeV and 106 GeV, accompanied by an opposite-charge muon pair. Events with a } Z \text{ candidate decaying to a pair of electrons or muons in the aforementioned mass range are excluded. Isolation and transverse impact parameter significance requirements are applied only to the leptons of the } e\mu \text{ pair.} \]

In order to estimate the reducible \( \ell\ell + ee \) background, a control region is formed by relaxing the selection criteria for the electrons of the sub-leading pair. The different sources of electron background are then separated into categories consisting of non-prompt leptons from heavy flavour decays, electrons from photon conversions and jets mis-identified as electrons, using appropriate discriminating variables [91]. This method allows the sum of the \( Z + \text{jets and } t\bar{t} \) background contributions to be estimated. As a cross-check, the same method is also applied to a similar control region containing same-charge sub-leading electron pairs. An additional cross-check of the \( \ell\ell + ee \) background estimation is performed by using a control region with same-charge sub-leading electron pairs, where the three highest \( p_T \) leptons satisfy all the analysis criteria whereas the selection cuts are relaxed for the remaining electrons. All the cross-checks yield consistent results.

The data-driven background estimates are summarised in Table 2. The distribution of \( m_{34\ell} \), for events selected by the analysis except that the isolation and transverse impact parameter requirements for the sub-leading lepton pair are removed, is presented in Fig. 1.

### 4.3. Systematic uncertainties

The uncertainties on the integrated luminosities are determined to be 1.8% for the 7 TeV data and 3.6% for the 8 TeV data using the techniques described in Ref. [92].

The uncertainties on the lepton reconstruction and identification efficiencies and on the momentum scale and resolution are determined using samples of \( W, Z \) and \( J/\psi \) decays [85, 84]. The relative uncertainty on the signal acceptance due to the uncertainty on the muon reconstruction and identification efficiency is \( \pm 0.7\% \) (\( \pm 0.5\% / \pm 0.5\% \)) for the 4\( \mu \) (2e2\( \mu \)/2\( \mu 2\mu \)) channel for \( m_{4\ell} = 600 \) GeV and increases to \( \pm 0.9\% \) (\( \pm 0.8\% / \pm 0.5\% \)) for \( m_{4\ell} = 115 \) GeV. Similarly, the relative uncertainty on the signal acceptance due to the uncertainty on the electron reconstruction and identification efficiency is \( \pm 2.6\% \) (\( \pm 1.7\% / \pm 1.8\% \)) for the 4\( e \) (2e2\( \mu \)/2\( \mu 2\mu \)) channel for \( m_{4\ell} = 600 \) GeV and reaches \( \pm 8.0\% \)
The impact of not using the theoretical constraints on the interpretation of the excess of events near $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data. The signal expectation for a SM Higgs with $m_H = 125$ GeV is also shown.

Table 3
The numbers of expected signal ($m_H = 125$ GeV) and background events, together with the numbers of observed events in the data, in a window of size $\pm 5$ GeV around 125 GeV, for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data.

<table>
<thead>
<tr>
<th>Signal</th>
<th>$ZZ^{(*)}$</th>
<th>$Z + \text{jets}$, tl</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4\mu$</td>
<td>2.09 ± 0.30</td>
<td>1.12 ± 0.05</td>
<td>0.13 ± 0.04</td>
</tr>
<tr>
<td>$2\ell2\mu/2\mu2\ell$</td>
<td>2.29 ± 0.33</td>
<td>0.80 ± 0.05</td>
<td>1.27 ± 0.19</td>
</tr>
<tr>
<td>$4\ell$</td>
<td>0.90 ± 0.14</td>
<td>0.44 ± 0.04</td>
<td>1.09 ± 0.20</td>
</tr>
</tbody>
</table>

The theoretical uncertainties associated with the signal are described in detail in Section 8. For the SM $ZZ^{(*)}$ background, which is estimated from MC simulation, the uncertainty on the total yield due to the QCD scale uncertainty is $\pm 5\%$, while the effect of the PDF and $m_{Z^*}$ uncertainties is $\pm 4\%$ (6%) for processes initiated by quarks (gluons) [53]. In addition, the dependence of these uncertainties on the four-lepton invariant mass spectrum has been taken into account as discussed in Ref. [53]. Though a small excess of events is observed for $m_{4\ell} > 160$ GeV, the measured $ZZ^{(*)} \rightarrow 4\ell$ cross section [93] is consistent with the SM theoretical prediction. The impact of not using the theoretical constraints on the $ZZ^{(*)}$ yield on the search for a Higgs boson with $m_H < 2m_Z$ has been studied in Ref. [87] and has been found to be negligible. The impact of the interference between a Higgs signal and the non-resonant $gg \rightarrow ZZ^{(*)}$ background is small and becomes negligible for $m_t < 2m_Z$ [94].

5. $H \rightarrow \gamma\gamma$ channel

The search for the SM Higgs boson through the decay $H \rightarrow \gamma\gamma$ is performed in the mass range between 110 GeV and 150 GeV. The dominant background is SM diphoton production ($\gamma\gamma$); contributions also come from $\gamma + \text{jet}$ and jet + jet production with one or two jets mis-identified as photons ($\gamma\gamma$) and from the Drell–Yan process. The 7 TeV data have been re-analysed and the results combined with those from the 8 TeV data. Among other changes to the analysis, a new category of events with two jets is introduced, which enhances the sensitivity to the VBF process. Higgs boson events produced by the VBF process have two forward jets, originating from the two scattered quarks, and tend to be devoid of jets in the central region. Overall, the sensitivity of the analysis has been improved by about 20% with respect to that described in Ref. [95].

5.1. Event selection

The data used in this channel are selected using a diphoton trigger [96], which requires two clusters formed from energy depositions in the electromagnetic calorimeter. An $E_T$ threshold of 20 GeV is applied to each cluster for the 7 TeV data, while for the 8 TeV data the thresholds are increased to 35 GeV on the leading (the highest $E_T$) cluster and to 25 GeV on the sub-leading (the next-highest $E_T$) cluster. In addition, loose criteria are applied to the shapes of the clusters to match the expectations for electromagnetic showers initiated by photons. The efficiency of the trigger is greater than 99% for events passing the final event selection.

Events are required to contain at least one reconstructed vertex with at least two associated tracks with $p_T > 0.4$ GeV, as well as two photon candidates. Photon candidates are reconstructed in the fiducial region $|\eta| < 2.37$, excluding the calorimeter barrel/end-cap transition region $1.37 < |\eta| < 1.52$. Photons that convert to electron–positron pairs in the ID material can have one or two reconstructed tracks matched to the clusters in the calorimeter. The photon reconstruction efficiency is about 97% for $E_T > 30$ GeV.

In order to account for energy losses upstream of the calorimeter and energy leakage outside of the cluster, MC simulation results are used to calibrate the energies of the photon candidates; there are separate calibrations for unconverted and converted
candidates. The calibration is refined by applying $\eta$-dependent correction factors, which are of the order of $\pm1\%$, determined from measured $Z \rightarrow e^+e^-$ events. The leading (sub-leading) photon candidate is required to have $E_T > 40$ GeV (30 GeV).

Photon candidates are required to pass identification criteria based on shower shapes in the electromagnetic calorimeter and on energy leakage into the hadronic calorimeter [97]. For the 7 TeV data, this information is combined in a neural network, tuned to achieve a similar jet rejection as the cut-based selection described in Ref. [95], but with higher photon efficiency. For the 8 TeV data, cut-based criteria are used to ensure reliable photon performance for recently-recorded data. This cut-based selection has been tuned to be robust against pile-up by relaxing requirements on shower shape criteria more susceptible to pile-up, and tightening others. The photon identification efficiencies, averaged over $\eta$, range from 85% to above 95% for the $E_T$ range under consideration.

To further suppress the jet background, an isolation requirement is applied. The isolation transverse energy is defined as the sum of the transverse energy of positive-energy topological clusters, as described in Section 4, within a cone of size $\Delta R = 0.4$ around the photon candidate, excluding the region within 0.125 $\times$ 0.175 in $\Delta \eta \times \Delta \phi$ around the photon barycentre. The distributions of the isolation transverse energy in data and simulation have been found to be in good agreement using electrons from $Z \rightarrow e^+e^-$ events and photons from $Z \rightarrow \ell^+\ell^-\gamma$ events. Remaining small differences are taken into account as a systematic uncertainty. Photon candidates are required to have an isolation transverse energy of less than 4 GeV.

5.2. Invariant mass reconstruction

The invariant mass of the two photons is evaluated using the photon energies measured in the calorimeter, the azimuthal angle $\phi$ between the photons as determined from the positions of the photons in the calorimeter, and the values of $\eta$ calculated from the position of the identified primary vertex and the impact points of the photons in the calorimeter.

The primary vertex of the hard interaction is identified by combining the following information in a global likelihood: the directions of flight of the photons as determined using the longitudinal segmentation of the electromagnetic calorimeter (calorimeter pointing), the parameters of the beam spot, and the $\sum p_T^2$ of the tracks associated with each reconstructed vertex. In addition, for the 7 TeV data analysis, the reconstructed conversion vertex is used in the likelihood for converted photons with tracks containing hits in the silicon layers of the ID. The calorimeter pointing is sufficient to ensure that the contribution of the opening angle between the photons to the mass resolution is negligible. Using the calorimeter pointing alone, the resolution of the vertex $z$ coordinate is $\sim 15$ mm, improving to $\sim 6$ mm for events with two reconstructed converted photons. The tracking information from the ID improves the identification of the vertex of the hard interaction, which is needed for the jet selection in the 2-jet category.

With the selection described in Section 5.1, in the diphoton invariant mass range between 100 GeV and 160 GeV, 23788 and 35251 diphoton candidates are observed in the 7 TeV and 8 TeV data samples, respectively.

Data-driven techniques [98] are used to estimate the numbers of $\gamma\gamma$, $\gamma j$ and $jj$ events in the selected sample. The contribution from the Drell-Yan background is determined from a sample of $Z \rightarrow e^+e^-$ decays in data where either one or both electrons pass the photon selection. The measured composition of the selected sample is approximately 74%, 22%, 3% and 1% for the $\gamma\gamma$, $\gamma j$, $jj$ and Drell-Yan processes, respectively, demonstrating the dominance of the irreducible diphoton production. This decomposition is not directly used in the signal search; however, it is used to study the parameterisation of the background modelling.

5.3. Event categorisation

To increase the sensitivity to a Higgs boson signal, the events are separated into ten mutually exclusive categories having different mass resolutions and signal-to-background ratios. An exclusive category of events containing two jets improves the sensitivity to VBF. The other nine categories are defined by the presence or not of converted photons, $\eta$ of the selected photons, and $p_T$, the component of the diphoton $p_T$ that is orthogonal to the axis defined by the difference between the two photon momenta [99,100].

Jets are reconstructed [101] using the anti-$k_T$ algorithm [102] with radius parameter $R = 0.4$. At least two jets with $|\eta| < 4.5$ and $p_T > 25$ GeV are required in the 2-jet selection. In the analysis of the 8 TeV data, the $p_T$ threshold is raised to 30 GeV for jets with $2.5 < |\eta| < 4.5$. For jets in the ID acceptance ($|\eta| < 2.5$), the fraction of the sum of the $p_T$ of tracks, associated with the jet and matched to the selected primary vertex, with respect to the sum of the $p_T$ of tracks associated with the jet (jet vertex fraction, JVF) is required to be at least 0.75. This requirement on the JVF reduces the number of jets from proton–proton interactions not associated with the primary vertex. Motivated by the VBF topology, three additional cuts are applied in the 2-jet selection: the difference of the pseudorapidity between the leading and sub-leading jets (tag jets) is required to be larger than 2.8. The invariant mass of the tag jets has to be larger than 400 GeV, and the azimuthal angle difference between the diphoton system and the system of the tag jets has to be larger than 2.6. About 70% of the signal events in the 2-jet category come from the VBF process.

The other nine categories are defined as follows: events with two unconverted photons are separated into unconverted central ($|\eta| < 0.75$ for both candidates) and unconverted rest (all other events), events with at least one converted photon are separated into converted central ($|\eta| < 0.75$ for both candidates), converted transition (at least one photon with $1.3 < |\eta| < 1.75$) and converted rest (all other events). Except for the converted transition category, each category is further divided by a cut at $p_T = 60$ GeV into two categories, low $p_T$ and high $p_T$. MC studies show that signal events, particularly those produced via VBF or associated production ($WH/ZH$ and $t\bar{t}H$), have on average larger $p_T$ than background events. The number of data events in each category, as well as the sum of all the categories, which is denoted inclusive, are given in Table 4.

5.4. Signal modelling

The description of the Higgs boson signal is obtained from MC, as described in Section 3. The cross sections multiplied by the branching ratio into two photons are given in Table 4 for $m_H = 126.5$ GeV. The number of signal events produced via the ggF process is rescaled to take into account the expected destructive interference between the $gg \rightarrow \gamma\gamma$ continuum background and ggF [103], leading to a reduction of the production rate by 2–5% depending on $m_H$ and the event category. For both the 7 TeV and 8 TeV MC samples, the fractions of ggF, VBF, $WH$, $ZH$ and $t\bar{t}H$ production are approximately 88%, 7%, 3%, 2% and 0.5%, respectively, for $m_H = 126.5$ GeV.

In the simulation, the shower shape distributions are shifted slightly to improve the agreement with the data [97], and the

\[ p_T = \frac{(p_T^1 + p_T^2) \times (p_T^2 - p_T^1)}{p_T^1 - p_T^2}, \] where $p_T^1$ and $p_T^2$ are the transverse momenta of the two photons.
Table 4

<table>
<thead>
<tr>
<th>Category</th>
<th>7 TeV</th>
<th>8 TeV</th>
<th>FWHM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\sigma \times (H \rightarrow \gamma \gamma)) [fb]</td>
<td>(N_b)</td>
<td>(N_s)</td>
</tr>
<tr>
<td>Unconv. central, low (p_T)</td>
<td>2054</td>
<td>10.5</td>
<td>2945</td>
</tr>
<tr>
<td>Unconv. central, high (p_T)</td>
<td>97</td>
<td>1.5</td>
<td>173</td>
</tr>
<tr>
<td>Unconv. rest, low (p_T)</td>
<td>7129</td>
<td>21.6</td>
<td>12136</td>
</tr>
<tr>
<td>Unconv. rest, high (p_T)</td>
<td>444</td>
<td>2.8</td>
<td>785</td>
</tr>
<tr>
<td>Conv. central, low (p_T)</td>
<td>1483</td>
<td>6.7</td>
<td>2015</td>
</tr>
<tr>
<td>Conv. central, high (p_T)</td>
<td>77</td>
<td>1.0</td>
<td>113</td>
</tr>
<tr>
<td>Conv. rest, low (p_T)</td>
<td>8313</td>
<td>21.1</td>
<td>11099</td>
</tr>
<tr>
<td>Conv. rest, high (p_T)</td>
<td>501</td>
<td>2.7</td>
<td>706</td>
</tr>
<tr>
<td>Conv. transition</td>
<td>3591</td>
<td>9.5</td>
<td>5140</td>
</tr>
<tr>
<td>2-jet</td>
<td>89</td>
<td>2.2</td>
<td>139</td>
</tr>
<tr>
<td>All categories (inclusive)</td>
<td>23788</td>
<td>79.6</td>
<td>35251</td>
</tr>
</tbody>
</table>

The photon energy resolution is broadened (by approximately 1% in the barrel calorimeter and 1.2–2.1% in the end-cap regions) to account for small differences observed between \(Z \rightarrow e^+e^-\) data and MC events. The signal yields expected for the 7 TeV and 8 TeV data samples are given in Table 4. The overall selection efficiency is about 40%.

The shape of the invariant mass of the signal in each category is modelled by the sum of a Crystal Ball function [104], describing the core of the distribution with a width \(\sigma_{CB}\), and a Gaussian contribution describing the tails (amounting to < 10%) of the mass distribution. The expected full-width-at-half-maximum (FWHM) is 3.9 GeV and \(\sigma_{CB}\) is 1.6 GeV for the inclusive sample. The resolution varies with event category (see Table 4); the FWHM is typically a factor 2.3 larger than \(\sigma_{CB}\).

5.5. Background modelling

The background in each category is estimated from data by fitting the diphoton mass spectrum in the mass range 100–160 GeV with a selected model with free parameters of shape and normalisation. Different models are chosen for the different categories to achieve a good compromise between limiting the size of a potential bias while retaining good statistical power. A fourth-order Bernstein polynomial function [105] is used for the unconverted rest (low \(p_T\)), converted rest (low \(p_T\)) and inclusive categories, an exponential function of a second-order polynomial for the unconverted central (low \(p_T\)), converted central (low \(p_T\)) and converted transition categories, and an exponential function for all others.

Studies to determine the potential bias have been performed using large samples of simulated background events complemented by data-driven estimates. The background shapes in the simulation have been cross-checked using data from control regions. The potential bias for a given model is estimated, separately for each category, by performing a maximum likelihood fit to large samples of simulated background events in the mass range 100–160 GeV, of the sum of a signal plus the given background model. The signal shape is taken to follow the expectation for a SM Higgs boson; the signal yield is a free parameter of the fit. The potential bias is defined by the largest absolute signal yield obtained from the likelihood fit to the simulated background samples for hypothesised Higgs boson masses in the range 110–150 GeV. A pre-selection of background parameterisations is made by requiring that the potential bias, as defined above, is less than 20% of the statistical uncertainty on the fitted signal yield. The pre-selected parameterisation in each category with the best expected sensitivity for \(m_H = 125\) GeV is selected as the background model.

The largest absolute signal yield as defined above is taken as the systematic uncertainty on the background model. It amounts to ±(0.2–4.6) and ±(0.3–6.8) events, depending on the category for the 7 TeV and 8 TeV data samples, respectively. In the final fit to the data (see Section 5.7) a signal-like term is included in the likelihood function for each category. This term incorporates the estimated potential bias, thus providing a conservative estimate of the uncertainty due to the background modelling.

5.6. Systematic uncertainties

Hereafter, in cases where two uncertainties are quoted, they refer to the 7 TeV and 8 TeV data, respectively. The dominant experimental uncertainty on the signal yield (±8%, ±11%) comes from the photon reconstruction and identification efficiency, which is estimated with data using electrons from \(Z\) decays and photons from \(Z \rightarrow e^+e^-\) events. Dipole modelling also affects the expected yields and contributes to the uncertainty (±4%). Further uncertainties on the signal yield are related to the trigger (±1%), photon isolation (±0.4%, ±0.5%) and luminosity (±1.8%, ±3.6%). Uncertainties due to the modelling of the underlying event are ±6% for VBF and ±30% for other production processes in the 2-jet category. Uncertainties on the predicted cross sections and branching ratios are summarised in Section 8.

The uncertainty on the expected fractions of signal events in each category is described in the following. The uncertainty on the knowledge of the material in front of the calorimeter is used to derive the amount of possible event migration between the converted and unconverted categories (±4%). The uncertainty from pile-up on the population of the converted and unconverted categories is ±2%. The uncertainty from pile-up calibration is ±6% (for the 8 TeV data) for the 2-jet category, estimated from \(Z\)-plus-jets events by comparing data and MC. Different PDFs and scale variations in the HQET calculations are used to derive possible event migration among categories (±9%) due to the modelling of the Higgs boson kinematics.

The total uncertainty on the mass resolution is ±14%. The dominant contribution (±12%) comes from the uncertainty on the energy resolution of the calorimeter, which is determined from \(Z \rightarrow e^+e^-\) events. Smaller contributions come from the imperfect knowledge of the material in front of the calorimeter, which affects the extrapolation of the calibration from electrons to photons (±6%), and from pile-up (±4%).

5.7. Results

The distributions of the invariant mass, \(m_{\gamma\gamma}\), of the diphoton events, summed over all categories, are shown in Fig. 4(a) and (b). The result of a fit including a signal component fixed to \(m_H = 126.5\) GeV and a background component described by a fourth-order Bernstein polynomial is superimposed.

The statistical analysis of the data employs an unbinned likelihood function constructed from those of the ten categories of the 7 TeV and 8 TeV data samples. To demonstrate the sensitivity of this likelihood analysis, Figs. 4(c) and (d) also show the mass spectrum obtained after weighting events with category-dependent factors reflecting the signal-to-background ratios. The weight \(w_i\) for events in category \(i \in [1, 10]\) for the 7 TeV and 8 TeV data samples is defined to be \(\ln(1 + S_i/B_i)\), where \(S_i\) is 90% of the expected signal for \(m_H = 126.5\) GeV, and \(B_i\) is the integral, in
6.1. Event selection

For the 8 TeV $H \rightarrow WW^{(*)} \rightarrow e\mu\nu\bar{\nu}$ search, the data are selected using inclusive single-muon and single-electron triggers. Both triggers require an isolated lepton with $p_T > 24$ GeV. Quality criteria are applied to suppress non-collision backgrounds such as cosmic-ray muons, beam-related backgrounds, and noise in the calorimeters. The primary vertex selection follows that described in Section 4. Candidates for the $H \rightarrow WW^{(*)} \rightarrow e\mu\nu\bar{\nu}$ search are pre-selected by requiring exactly two opposite-charge leptons of different flavours, with $p_T$ thresholds of 25 GeV for the leading lepton and 15 GeV for the sub-leading lepton. Events are classified into two exclusive lepton channels depending on the flavour of the leading lepton, where $e\mu$ ($\mu\mu$) refers to events with a leading electron (muon). The dilepton invariant mass is required to be greater than 10 GeV.

The lepton selection and isolation have more stringent requirements than those used for the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ analysis (see Section 4), to reduce the larger background from non-prompt leptons in the $\ell\nu\ell\nu$ final state. Electron candidates are selected using a combination of tracking and calorimetric information [85]; the criteria are optimised for background rejection, at the expense of some reduced efficiency. Muon candidates are restricted to those with matching MS and ID tracks [84], and therefore are reconstructed over $|\eta| < 2.5$. The isolation criteria require the scalar sums of the $p_T$ of charged particles and of calorimeter topological clusters within $\Delta R = 0.3$ of the lepton direction (excluding the lepton itself) each to be less than 0.12–0.20 times the lepton $p_T$. The exact value differs between the criteria for tracks and calorimeter clusters, for both electrons and muons, and depends on the lepton $p_T$. Jet selections follow those described in Section 5.3, except that the JVF is required to be greater than 0.5.

Since two neutrinos are present in the signal final state, events are required to have large $E_T^{miss}$. $E_T^{miss}$ is the negative vector sum of the transverse momenta of the reconstructed objects, including muons, electrons, photons, jets, and clusters of calorimeter cells not associated with these objects. The quantity $E_T^{miss}$ used in this analysis is required to be greater than 25 GeV and is defined as:

$$E_T^{miss} = E_T^{miss} \sin \Delta \phi_{min}, \quad \Delta \phi_{min} = \text{min}(\Delta \phi, \pi)$$

and $E_T^{miss}$ is the magnitude of the vector $E_T^{miss}$. Here, $\Delta \phi$ is the angle between $E_T^{miss}$ and the transverse momentum of the nearest lepton or jet with $p_T > 25$ GeV. Compared to $E_T^{miss}$, $E_T^{miss}$ has increased rejection power for events in which the $E_T^{miss}$ is generated by a neutrino in a jet or the mismeasurement of an object, since in those events the $E_T^{miss}$ tends to point in the direction of the object. After the lepton isolation and $E_T^{miss}$ requirements that define the pre-selected sample, the multitjet background is negligible and the Drell–Yan background is much reduced. The Drell–Yan contribution becomes very small after the topological selections, described below, are applied.

The background rate and composition depend significantly on the jet multiplicity, as does the signal topology. Without accompanying jets, the signal originates almost entirely from the ggF process and the background is dominated by $WW$ events. In contrast, when produced in association with two or more jets, the signal contains a much larger contribution from the VBF process compared to the ggF process, and the background is dominated by $tt$ production. Therefore, to maximise the sensitivity to SM Higgs events, further selection criteria depending on the jet multiplicity are applied to the pre-selected sample. The data are subdivided into 0-jet, 1-jet and 2-jet search channels according to the number of jets present.

A window containing $S_i$, of a background-only fit to the data. The values $S_i/B_i$ have only a mild dependence on $m_{H}$.

The statistical interpretation of the excess of events near $m_{\gamma\gamma} = 126.5$ GeV in Fig. 4 is presented in Section 9.
of jets in the final state, with the 2-jet channel also including higher jet multiplicities.

Owing to spin correlations in the $W W^{(*)}$ system arising from the spin-0 nature of the SM Higgs boson and the V-A structure of the $W$ boson decay vertex, the charged leptons tend to emerge from the primary vertex pointing in the same direction [107]. This kinematic feature is exploited for all jet multiplicities by requiring that $|\Delta \phi_{\ell \ell}| < 1.8$, and the dilepton invariant mass, $m_{\ell\ell}$, be less than 50 GeV for the 0-jet and 1-jet channels. For the 2-jet channel, the $m_{\ell\ell}$ upper bound is increased to 80 GeV.

In the 0-jet channel, the magnitude $p_T^{\ell\ell}$ of the transverse momentum of the dilepton system, $p_T^{\ell\ell} = p_T^{\ell1} + p_T^{\ell2}$, is required to be greater than 30 GeV. This improves the rejection of the Drell–Yan background.

In the 1-jet channel, backgrounds from top quark production are suppressed by rejecting events containing a $b$-tagged jet, as determined using a $b$-tagging algorithm that uses a neural network and exploits the topology of weak decays of $b$- and $c$-hadrons [108]. The total transverse momentum, $p_T^{\ell\ell}$, defined as the magnitude of the vector sum $p_T^{\ell\ell} = p_T^{\ell1} + p_T^{\ell2} + p_T^{miss} T$, is required to be smaller than 30 GeV to suppress top background events that have jets with $p_T$ below the threshold defined for jet counting. In order to reject the background from $Z \rightarrow \tau\tau$, the $\tau\tau$ invariant mass, $m_{\tau\tau}$, is computed under the assumptions that the reconstructed leptons are $\tau$ lepton decay products. In addition the neutrinos produced in these decays are assumed to be the only source of $E^{miss}$ and to be collinear with the leptons [109]. Events with $|m_{\tau\tau} - m_Z| < 25$ GeV are rejected if the collinear approximation yields a physical solution.

The 2-jet selection follows the 1-jet selection described above, with the $p_T^{\ell\ell}$ definition modified to include all selected jets. Motivated by the VBF topology, several additional criteria are applied to the tag jets, defined as the two highest-$p_T$ jets in the event. These are required to be separated in rapidity by a distance $|\Delta y_{j1}| > 3.8$ and to have an invariant mass, $m_{jj}$, larger than 500 GeV. Events with an additional jet with $p_T > 20$ GeV between the tag jets ($y_{j1} < y < y_{j2}$) are rejected.

A transverse mass variable, $m_T$ [110], is used to test for the presence of a signal for all jet multiplicities. This variable is defined as:

$$m_T = \sqrt{(E_T^{miss} + E_T^{\ell\ell})^2 - |p_T^{\ell\ell} + E_T^{miss}|^2},$$

where $E_T^{\ell\ell} = \sqrt{p_T^{\ell1} + p_T^{\ell2} + 2p_T^{\ell1}p_T^{\ell2}\cos(\phi^{\ell1} - \phi^{\ell2})}$. The statistical analysis of the data uses a fit to the $m_T$ distribution in the signal region after the $\Delta \phi_{\ell\ell}$ requirement (see Section 6.4), which results in increased sensitivity compared to the analysis described in Ref. [111].

For a SM Higgs boson with $m_H = 125$ GeV, the cross section times branching ratio to the $e\mu\nu\nu$ final state is 88 fb for $\sqrt{s} = 7$ TeV, increasing to 112 fb at $\sqrt{s} = 8$ TeV. The combined acceptance times efficiency of the $8$ TeV 0-jet and 1-jet selection relative to the ggF production cross section times branching ratio is about 7.4%. The acceptance times efficiency of the $8$ TeV 2-jet selection relative to the VBF production cross section times branching ratio is about 14%. Both of these figures are based on the number of events selected before the final $m_T$ criterion is applied (as described in Section 6.4).

6.2. Background normalisation and control samples

The leading backgrounds from SM processes producing two isolated high-$p_T$ leptons are $W W$ and top (in this section, "top" background always includes both $t\bar{t}$ and single top, unless otherwise noted). These are estimated using partially data-driven techniques based on normalising the MC predictions to the data in control regions dominated by the relevant background source. The $W +$ jets background is estimated from data for all jet multiplicities. Only the small backgrounds from Drell–Yan and diboson processes other than $W W$, as well as the $W W$ background for the 2-jet analysis, are estimated using MC simulation.

The control and validation regions are defined by selections similar to those used for the signal region but with some criteria reversed or modified to obtain signal-depleted samples enriched in a particular background. The term “validation region” distinguishes these regions from the control regions that are used to directly normalise the backgrounds. Some control regions have significant contributions from backgrounds other than the targeted one, which introduces dependencies among the background estimates. These correlations are fully incorporated in the fit to the $m_T$ distribution. In the following sections, each background estimate is described after any others on which it depends. Hence, the largest background ($W W$) is described last.

6.2.1. $W +$ jets background estimation

The $W +$ jets background contribution is estimated using a control sample of events where one of the two leptons satisfies the identification and isolation criteria described in Section 6.1, and the other lepton fails these criteria but satisfies a loosened selection (denoted “anti-identified”). Otherwise, events in this sample are required to pass all the signal selections. The dominant contribution to this sample comes from $W +$ jets events in which a jet produces an object that is reconstructed as a lepton. This object may be either a true electron or muon from the decay of a heavy quark, or else a product of the fragmentation identified as a lepton candidate.

The contamination in the signal region is obtained by scaling the number of events in the data control sample by a transfer factor. The transfer factor is defined here as the ratio of the number of identified lepton candidates passing all selections to the number of anti-identified leptons. It is calculated as a function of the anti-identified lepton $p_T$ using a data sample dominated by QCD jet production (dijet sample) after subtracting the residual contributions from leptons produced by leptonic $W$ and $Z$ decays, as estimated from data. The small remaining lepton contamination, which includes $W\gamma(\rightarrow e\gamma)$ events, is subtracted using MC simulation.

The processes producing the majority of same-charge dilepton events, $W +$ jets, $W\gamma(\rightarrow e\gamma)$, $W Z^{(*)}$, and $Z^{(*)}Z^{(*)}$, are all backgrounds in the opposite-charge signal region. $W +$ jets and $W\gamma(\rightarrow e\gamma)$ backgrounds are particularly important in a search optimised for a low Higgs boson mass hypothesis. Therefore, the normalisation and kinematic features of same-charge dilepton events are used to validate the predictions of these backgrounds. The predicted number of same-charge events after the $E_{T}^{miss}$ and zero-jet requirements is 216 ± 7 (stat) ± 42 (syst), while 182 events are observed in the data. Satisfactory agreement between data and simulation is observed in various kinematic distributions, including those of $\Delta \phi_{\ell\ell}$ (see Fig. 5(a)) and the transverse mass.

6.2.2. Top control sample

In the 0-jet channel, the top quark background prediction is first normalised using events satisfying the pre-selection criteria described in Section 6.1. This sample is selected without jet multiplicity or $b$-tagging requirements, and the majority of events contain top quarks. Non-top contributions are subtracted using predictions from simulation, except for $W +$ jets, which is estimated using data. After this normalisation is performed, the fraction of events with zero jets that pass all selections is evaluated. This fraction is small (about 3%), since the top quark decay $t \rightarrow Wb$
The resulting correction factor to a purely MC-based background estimate after all selections amounts to 1.11 ± 0.05 (stat). The Δφ_{jj}, Δφ_{jℓℓ}, and m_{T} requirements are included in the definition of the 2-jet control region. The resulting samples are dominated by top quark events. The small contributions from other sources are taken into account using MC simulation and the data-driven W + jets estimate. Good agreement between data and MC simulation is observed for the total numbers of events and the shapes of the m_{T} distributions. The resulting normalisation factors are 1.11 ± 0.05 for the 1-jet control region and 1.01 ± 0.26 for the 2-jet control region. Only the statistical uncertainties are quoted.

6.2.3. W ± W control sample

The MC predictions of the W ± W background in the 0-jet and 1-jet analyses, summed over lepton flavours, are normalised using control regions defined with the same selections as for the signal region except that the Δφ_{jj} requirement is removed and the upper bound on m_{T} is replaced with a lower bound: m_{T} > 80 GeV. The numbers of events and the shape of the m_{T} distribution in the control regions are in good agreement between data and MC, as shown in Fig. 1. W ± W production contributes about 70% of the events in the 0-jet control region and about 45% in the 1-jet region. Contaminations from sources other than W ± W are derived as for the signal region, including the data-driven W + jets and top estimates. The resulting normalisation factors with their associated statistical uncertainties are 1.06 ± 0.06 for the 0-jet control region and 0.99 ± 0.15 for the 1-jet control region.

6.3. Systematic uncertainties

The systematic uncertainties that have the largest impact on the sensitivity of the search to the theoretical uncertainties associated with the signal. These are described in Section 9. The main experimental uncertainties are associated with the JES, the jet energy resolution (JER), pile-up, E_{miss}^{T} Δσ_{ℓℓ}, and the b-tagging efficiency, the W + jets transfer factor, and the integrated luminosity. The largest uncertainties on the backgrounds include W ± W normalisation and modelling, top normalisation, and W ± γ(∗) normalisation. The 2-jet systematic uncertainties are dominated by the statistical uncertainties in the data and the MC simulation, and are therefore not discussed further.

Variations of the jet energy scale within the systematic uncertainties can cause events to migrate between the jet bins. The uncertainty on the JES varies from ±2% to ±9% as a function of jet p_{T} and |η| for jets with p_{T} > 25 GeV and |η| < 4.5 [101]. The largest impact of this uncertainty on the total signal (background) yield amounts to 7% (4%) in the 0-jet (1-jet) bin. The uncertainty on the JER is estimated from in situ measurements and it impacts mostly the 1-jet channel, where its effect on the total signal and background yields is 4% and 2%, respectively. An additional contribution to the JES uncertainty arises from pile-up, and is estimated to vary between ±1% and ±5% for multiple pp collisions in the same bunch crossing and up to ±10% for neighbouring bunch crossings. This uncertainty affects mainly the 1-jet channel, where its impact on the signal and background yields is 4% and 2%, respectively. An additional contribution to the JES uncertainty arises from pile-up, and is estimated to vary between ±1% and ±5% for multiple pp collisions in the same bunch crossing and up to ±10% for neighbouring bunch crossings. This uncertainty affects mainly the 1-jet channel, where its impact on the total signal and background yields is 4% and 2%, respectively. An additional contribution to the JES uncertainty arises from pile-up, and is estimated to vary between ±1% and ±5% for multiple pp collisions in the same bunch crossing and up to ±10% for neighbouring bunch crossings. This uncertainty affects mainly the 1-jet channel, where its impact on the total signal and background yields is 4% and 2%, respectively.
The expected numbers of signal ($m_H = 125$ GeV) and background events after all selections, including a cut on the transverse mass of $0.75 m_H < m_T < m_H$ for $m_H = 125$ GeV. The observed numbers of events in data are also displayed. The $e\mu$ and $\mu\mu$ channels are combined. The uncertainties shown are the combination of the statistical and all systematic uncertainties, taking into account the constraints from control samples. For the 2-jet analysis, backgrounds with fewer than 0.01 expected events are marked with ‘–’.

<table>
<thead>
<tr>
<th>mH [GeV]</th>
<th>0-jet</th>
<th>1-jet</th>
<th>2-jet</th>
</tr>
</thead>
<tbody>
<tr>
<td>101 ± 13</td>
<td>12 ± 3</td>
<td>1.9 ± 1.1</td>
<td>0.10 ± 0.10</td>
</tr>
<tr>
<td>8 ± 2</td>
<td>6 ± 2</td>
<td>0.15 ± 0.10</td>
<td></td>
</tr>
<tr>
<td>7 ± 7</td>
<td>2 ± 1</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Total signal</td>
<td>20 ± 4</td>
<td>5 ± 2</td>
<td>0.34 ± 0.07</td>
</tr>
<tr>
<td>Total background</td>
<td>142 ± 16</td>
<td>26 ± 6</td>
<td>0.35 ± 0.18</td>
</tr>
<tr>
<td>Observed</td>
<td>185</td>
<td>38</td>
<td>0</td>
</tr>
</tbody>
</table>

The potential impact of interference between resonant (Higgs-mediated) and non-resonant $gg \rightarrow WW$ diagrams [116] for $m_T > m_H$ was investigated and found to be negligible. The effect of the $WW$ normalisation, modelling, and shape systematics on the total background yield is 9% for the 0-jet channel and 19% for the 1-jet channel. The uncertainty on the shape of the total background is dominated by the uncertainties on the normalisations of the individual backgrounds. The main uncertainties on the top background in the 0-jet analysis include those associated with interference effects between $tt$ and single top, initial state an final state radiation, $b$-tagging, and JER. The impact on the total background yield in the 0-jet bin is 3%. For the 1-jet analysis, the impact of the top background on the total yield is 14%. Theoretical uncertainties on the $W\gamma$ background normalisation are evaluated for each jet bin using the procedure described in Ref. [117]. They are ±11% for the 0-jet bin and ±50% for the 1-jet bin. For $W\gamma^*$ with $m_{T\gamma} < 7$ GeV, a $k$-factor of 1.3 ± 0.3 is applied to the MadGraph LO prediction based on the comparison with the MC@NLO calculation. The $k$-factor for $W\gamma^*/WZ^{(*)}$ with $m_{T\gamma} > 7$ GeV is 1.5 ± 0.5. Theses uncertainties affect mostly the 1-jet channel, where their impact on the total background yield is approximately 4%.

6.4. Results

Table 5 shows the numbers of events expected from a SM Higgs boson with $m_H = 125$ GeV and from the backgrounds, as well as the numbers of candidates observed in data, after application of all selection criteria plus an additional cut on $m_T$ of $0.75 m_H < m_T < m_H$. The uncertainties shown in Table 5 include the systematic uncertainties discussed in Section 6.3, constrained by the use of the control regions discussed in Section 6.2. An excess of events relative to the background expectation is observed in the data.

Fig. 6 shows the distribution of the transverse mass after all selection criteria in the 0-jet and 1-jet channels combined, and for both lepton channels together.

The statistical analysis of the data employs a binned likelihood function constructed as the product of Poisson probability terms for the $e\mu$ channel and the $\mu\mu$ channel. The mass-dependent cuts on $m_T$ described above are not used. Instead, the 0-jet (1-jet) signal regions are subdivided into five (three) $m_T$ bins. For the 2-jet signal region, only the results integrated over $m_T$ are used, due to the small number of events in the final sample. The statistical interpretation of the observed excess of events is presented in Section 9.

7. Statistical procedure

The statistical procedure used to interpret the data is described in Refs. [17,118–121]. The parameter of interest is the global signal strength factor $\mu$, which acts as a scale factor on the total number of events predicted by the Standard Model for the Higgs boson signal. This factor is defined such that $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ corresponds to the SM Higgs boson signal in addition to the background. Hypothesised values of $\mu$ are tested with a statistic $\lambda(\mu)$ based on the profile likelihood ratio [122]. This test statistic extracts the information on the signal strength from a full likelihood fit to the data. The likelihood function includes all the parameters that describe the systematic uncertainties and their correlations.

Exclusion limits are based on the $CL_s$ prescription [123]; a value of $\mu$ is regarded as excluded at 95% CL when $CL_s$ is less than 5%. A SM Higgs boson with mass $m_H$ is considered excluded at 95% confidence level (CL) when $\mu = 1$ is excluded at that mass. The significance of an excess in the data is first quantified with the local significance $\mu_s$ and then with the method described in Ref. [124]. The ratio of the global to the local probabilities, the trials factor used to correct for the “look elsewhere” effect, increases with the range of Higgs boson mass hypotheses considered, the mass resolutions of the channels involved in the combination, and the significance of the excess.

The statistical tests are performed in steps of values of the hypothesised Higgs boson mass $m_H$. The asymptotic approximation [122] upon which the results are based has been validated with the method described in Ref. [17].

The combination of individual search sub-channels for a specific Higgs boson decay, and the full combination of all search channels, are based on the global signal strength factor $\mu$ and on the identification of the nuisance parameters that correspond to the correlated sources of systematic uncertainty described in Section 8.

8. Correlated systematic uncertainties

The individual search channels that enter the combination are summarised in Table 6.
The main uncorrelated systematic uncertainties are described in Sections 4–6 for the $H \to ZZ^{(*)} \to 4\ell$, $H \to \gamma\gamma$ and $H \to WW^{(*)} \to \ell\nu\ell\nu$ channels and in Ref. [17] for the other channels. They include the background normalisations or background model parameters from control regions or sidebands, the Monte Carlo simulation statistical uncertainties and the theoretical uncertainties affecting the background processes.

The main sources of correlated systematic uncertainties are the following.

1. Integrated luminosity: The uncertainty on the integrated luminosity is considered as fully correlated among channels and amounts to ±3.9% for the 7 TeV data [132,133], except for the $H \to ZZ^{(*)} \to 4\ell$ and $H \to \gamma\gamma$ channels which were re-analysed; the uncertainty is ±1.8% [92] for these channels. The uncertainty is ±3.6% for the 8 TeV data.

2. Electron and photon trigger identification: The uncertainties in the trigger and identification efficiencies are treated as fully correlated for electrons and photons.

3. Electron and photon energy scales: The electron and photon energy scales in the $H \to ZZ^{(*)} \to 4\ell$ and $H \to \gamma\gamma$ channels are described by five parameters, which provide a detailed account of the sources of systematic uncertainty. They are related to the calibration method, the presampler energy scale in the barrel and end-cap calorimeters, and the material description upstream of the calorimeters.

4. Muon reconstruction: The uncertainties affecting muons are separated into those related to the ID and MS, in order to obtain a better description of the correlated effects among channels using different muon identification criteria and different ranges of muon $p_T$.

5. Jet energy scale and missing transverse energy: The jet energy scale and jet energy resolution are affected by uncertainties which depend on the $p_T$, $\eta$, and flavour of the jet. A simplified scheme is used in which independent JES and JER nuisance parameters are associated with final states with significantly different kinematic selections and sensitivity to scattering processes with different kinematic distributions or flavour composition. This scheme includes a specific treatment for b-jets. The sensitivity of the results to various assumptions about the correlation of these sources of uncertainty has been found to be negligible. An uncorrelated component of the uncertainty on $E_T^{miss}$ is included, in addition to the JES uncertainty, which is due to low energy jet activity not associated with reconstructed physics objects.

6. Theory uncertainties: Correlated theoretical uncertainties affect mostly the signal predictions. The QCD scale uncertainties for $m_H = 125$ GeV amount to ±4% for the ggF process, ±1% for the VBF and $WH/ZH$ processes, and ±4% for the t$t\bar{t}H$ process [52,53]; the small dependence of these uncertainties on $m_H$ is taken into account. The uncertainties on the predicted branching ratios amount to ±5%. The uncertainties related to the parton distribution functions amount to ±8% for the predominantly gluon-initiated ggF and $t\bar{t}H$ processes, and ±4% for the predominantly quark-initiated VBF and $WH/ZH$ processes [78,134–136]. The theoretical uncertainty associated with the exclusive Higgs boson production process with additional jets in the $H \to \gamma\gamma$, $H \to WW^{(*)} \to \ell\nu\ell\nu$ and $H \to \tau^+\tau^-$ channels is estimated using the prescription of Refs. [53,117,118], with the noticeable difference that an explicit calculation of the gluon-fusion process at NLO using MCFM [137] in the 2-jet category reduces the uncertainty on this non-negligible contribution to 25%. An additional theoretical uncertainty on the signal normalisation of ±150% ($m_H$/TeV)$^3$ (e.g. ±4% for $m_H = 300$ GeV) accounts for effects related to off-shell Higgs boson production and interference with other SM processes [53].

Sources of systematic uncertainty that affect both the 7 TeV and the 8 TeV data are taken as fully correlated. The uncertainties on background estimates based on control samples in the data are considered uncorrelated between the 7 TeV and 8 TeV data.

9. Results

The addition of the 8 TeV data for the $H \to ZZ^{(*)} \to 4\ell$, $H \to \gamma\gamma$ and $H \to WW^{(*)} \to \ell\nu\ell\nu$ channels, as well as the improvements to the analyses of the 7 TeV data in the first two of these channels, bring a significant gain in sensitivity in the low-mass region with respect to the previous combined search [17].

9.1. Excluded mass regions

The combined 95% CL exclusion limits on the production of the SM Higgs boson, expressed in terms of the signal strength parameter $\mu$, are shown in Fig. 7(a) as a function of $m_H$. The expected 95% CL exclusion region covers the $m_H$ range from 110 GeV to
582 GeV. The observed 95% CL exclusion regions are 111–122 GeV and 131–559 GeV. Three mass regions are excluded at 99% CL, 113–114, 117–121 and 132–527 GeV, while the expected exclusion range at 99% CL is 113–532 GeV.

9.2. Observation of an excess of events

An excess of events is observed near $m_H = 126$ GeV in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels, both of which provide fully reconstructed candidates with high resolution in invariant mass, as shown in Figs. 8(a) and 8(b). These excesses are confirmed by the highly sensitive but low-resolution $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channel, as shown in Fig. 8(c).

The observed local $p_0$ values from the combination of channels, using the asymptotic approximation, are shown as a function of $m_H$ in Fig. 7(b) for the full mass range and in Fig. 9 for the low mass range.

The largest local significance for the combination of the 7 and 8 TeV data is found for a SM Higgs boson mass hypothesis of $m_H = 126.5$ GeV, where it reaches 6.0$\sigma$, with an expected value in the presence of a SM Higgs boson signal at that mass of 4.9$\sigma$ (see also Table 7). For the 2012 data alone, the maximum local significance for the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels combined is 4.9$\sigma$, and occurs at $m_H = 126.5$ GeV (3.8$\sigma$ expected).

The significance of the excess is mildly sensitive to uncertainties in the energy resolutions and energy scale systematic uncertainties for photons and electrons; the effect of the muon energy scale systematic uncertainties is negligible. The presence of these
uncertainties, evaluated as described in Ref. [138], reduces the local significance to 5.9σ.

The global significance of a local 5.9σ excess anywhere in the mass range 110–600 GeV is estimated to be approximately 5.1σ, increasing to 5.3σ in the range 110–150 GeV, which is approximately the mass range not excluded at the 99% CL by the LHC combined SM Higgs boson search [139] and the indirect constraints from the global fit to precision electroweak measurements [12].

9.3. Characterising the excess

The mass of the observed new particle is estimated using the profile likelihood ratio $\lambda(m_H)$ for $H \to ZZ^{(*)}$, $H \to 4\ell$ and $H \to gg$, the two channels with the highest mass resolution. The signal strength is allowed to vary independently in the two channels, although the result is essentially unchanged when restricted to the SM hypothesis $\mu = 1$. The leading sources of systematic uncertainty come from the electron and photon energy scales and resolutions. The resulting estimate for the mass of the observed particle is $m_{H} = 126.0 \pm 0.4$ (stat) $\pm 0.4$ (sys) GeV. The best-fit signal strength $\mu$ is shown in Fig. 7(c) as a function of $m_{H}$. The observed excess corresponds to $\mu = 1.4 \pm 0.3$ for $m_{H} = 126$ GeV, which is consistent with the SM Higgs boson hypothesis $\mu = 1$. A summary of the individual and combined best-fit values of the strength parameter for a SM Higgs boson mass hypothesis of 126 GeV is shown in Fig. 10, while more information about the three main channels is provided in Table 7.

In order to test which values of the strength and mass of a signal hypothesis are simultaneously consistent with the data, the profile likelihood ratio $\lambda(\mu, m_H)$ is used. In the presence of a strong signal, it will produce closed contours around the best-fit point $(\hat{\mu}, \hat{m}_H)$, while in the absence of a signal the contours will be upper limits on $\mu$ for all values of $m_H$.

Asymptotically, the test statistic $-2\ln \lambda(\mu, m_H)$ is distributed as a $\chi^2$ distribution with two degrees of freedom. The resulting 68% and 95% CL contours for the $H \to \gamma\gamma$ and $H \to WW^{(*)}$, $H \to \ell\ell\nu\nu$ channels are shown in Fig. 11, where the asymptotic approximations have been validated with ensembles of pseudo-experiments. Similar contours for the $H \to ZZ^{(*)}$, $H \to 4\ell$ channel are also shown in Fig. 11, although they are only approximate confidence intervals due to the smaller number of candidates in this channel. These contours in the $(\mu, m_H)$ plane take into account uncertainties in the energy scale and resolution.

The probability for a single Higgs boson-like particle to produce resonant mass peaks in the $H \to ZZ^{(*)}$, $H \to 4\ell$ and $H \to \gamma\gamma$ channels separated by more than the observed mass difference, allowing the signal strengths to vary independently, is about 8%.

The contributions from the different production modes in the $H \to \gamma\gamma$ channel have been studied in order to assess any tension between the data and the ratios of the production cross

![Fig. 10. Measurements of the signal strength parameter $\mu$ for $m_H = 126$ GeV for the individual channels and their combination.](image)

![Fig. 11. Confidence intervals in the $(\mu, m_H)$ plane for the $H \to ZZ^{(*)}$, $H \to \gamma\gamma$, and $H \to WW^{(*)}$, $H \to \ell\ell\nu\nu$ channels, including all systematic uncertainties. The markers indicate the maximum likelihood estimates $(\hat{\mu}, \hat{m}_H)$ in the corresponding channels (the maximum likelihood estimates for $H \to ZZ^{(*)}$, $H \to 4\ell$ and $H \to WW^{(*)}$, $H \to \ell\ell\nu\nu$ coincide).](image)
sections predicted in the Standard Model. A new signal strength parameter $\mu_i$ is introduced for each production mode, defined by $\mu_i = \sigma_i/\sigma_i^{SM}$. In order to determine the values of $(\mu_i, \mu_j)$ that are simultaneously consistent with the data, the profile likelihood ratio $\lambda(\mu_i, \mu_j)$ is used with the measured mass treated as a nuisance parameter.

Since there are four Higgs boson production modes at the LHC, two-dimensional contours require either some $\mu_i$ to be fixed, or multiple $\mu_i$ to be related in some way. Here, $\mu_{ggF}$ and $\mu_{VBF}$ have been grouped together as they scale with the $t\bar{t}H$ coupling in the SM, and are denoted by the common parameter $\mu_{ggF+VBF}$. Similarly, $\mu_{VBF}$ and $\mu_{VH}$ have been grouped together as they scale with the $WWH/ZZH$ coupling in the SM, and are denoted by the common parameter $\mu_{VBF+VH}$. Since the distribution of signal events among the 10 categories of the $H \to \gamma\gamma$ search is sensitive to these factors, constraints in the plane of $\mu_{ggF+VBF} \times B/BSM$ and $\mu_{VBF+VH} \times B/BSM$, where $B$ is the branching ratio for $H \to \gamma\gamma$, can be obtained (Fig. 12). Theoretical uncertainties are included so that the consistency with the SM expectation can be quantified. The data are compatible with the SM expectation at the 1.5$\sigma$ level.

10. Conclusion

Searches for the Standard Model Higgs boson have been performed in the $H \to ZZ^{(*)} \rightarrow 4\ell$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ channels with the ATLAS experiment at the LHC using 5.8–5.9 fb$^{-1}$ of $pp$ collision data recorded during April to June 2012 at a centre-of-mass energy of 8 TeV. These results are combined with earlier results [17], which are based on an integrated luminosity of 4.6–4.8 fb$^{-1}$ recorded in 2011 at a centre-of-mass energy of 7 TeV, except for the $H \to ZZ^{(*)} \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels, which have been updated with the improved analyses presented here.

The Standard Model Higgs boson is excluded at 95$\%$ CL in the mass range 111–559 GeV, except for the narrow region 122–131 GeV. In this region, an excess of events with significance 5.9$\sigma$, corresponding to $p_0 = 1.7 \times 10^{-9}$, is observed. The excess is driven by the two channels with the highest mass resolution, $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$, and the equally sensitive but low-resolution $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ channel. Taking into account the entire mass range of the search, 110–600 GeV, the global significance of the excess is 5.1$\sigma$, which corresponds to $p_0 = 1.7 \times 10^{-7}$.

These results provide conclusive evidence for the discovery of a new particle with mass 126.0$\pm$0.4 (stat)$\pm$0.4 (syst) GeV. The signal strength parameter $\mu$ has the value 1.4$\pm$0.3 at the fitted mass, which is consistent with the SM Higgs boson hypothesis $\mu = 1$. The decays to pairs of vector bosons whose net electric charge is zero identify the new particle as a neutral boson. The observation in the diphoton channel disfavours the spin-1 hypothesis [140, 141]. Although these results are compatible with the hypothesis that the new particle is the Standard Model Higgs boson, more data are needed to assess its nature in detail.

Acknowledgements

The results reported in this Letter would not have been possible without the outstanding performance of the LHC. We warmly thank CERN and the entire LHC exploitation team, including the operation, technical and infrastructure groups, and all the people who have contributed to the conception, design and construction of this superb accelerator. We thank also the support staff at our institutions without whose excellent contributions ATLAS could not have been successfully constructed or operated so efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET and ERC, European Union; IN2P3-CNRS, CEA-DSM/Irfu, France; GNOS, Georgia; BMBF, DFG, MPG and AvH Foundation, Germany; GSRT, Greece; ISF, MINERVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNISW, Poland; GRICES and FCT, Portugal; MERYS (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MVZT, Slovenia; DST/NRF, South Africa; MICINN, Spain; and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open access

This article is published Open Access at sciencedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

References

[12] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group,